

M.A.Aziz

aziz.muh@northeastern.edu. Boston, MA . [Website](#)

EDUCATION

Northeastern University

Jan 2021 - Dec 2025

PhD in Computer Science

LUMS

Sep 2016 - May 2020

BS in Computer Science

WORK EXPERIENCE

Privacy Researcher | Northeastern University

Jan 2021 – Present

- Led two projects that resulted in publications in top conferences:
 - I contributed to building a tool in **JavaScript** to assess whether websites comply with the **California Consumer Privacy Act (CCPA)**. Found **8%** adopting the IAB framework and **52%** of third parties inconsistently shared consent. Paper will be published at **PETS 2024**. [Code](#)
 - I built data pipelines in **Spark** to study **713** users' browsing habits and carried out data analysis. Discovered Google dominates **24** out of 30 online market segments. The [paper](#) got an **honorable mention award** at **AAI ICWSM** out of **138** papers presented at the conference.

Privacy Software Engineer | National Internet Observatory

Feb 2024 – Apr 2024

- Proposed novel ways to gather privacy-related data such as about **Google's Privacy Sandbox**.
- Added features in a web extension to collect data about cookies and the Topics API from **2000** users.
- Collaborated with the Ethics team to ensure that the data collected **respected user consent and privacy**.

Analyst Software Engineer | Afiniti

Jun 2020 – Sep 2020

- Analyzed and reported client's monthly performance based on revenue and sales using **SQL**.
- Identified and solved data issues that led to improved customer and agent model prediction by **4%**.

Research Assistant | LUMS

Sep 2019 – May 2020

- Built automated tools in Python and Javascript to extract mobile application data from Google Play Store.
- Uncovered a large number of Android applications (900) using Pay Per Install services potentially involved in inflating ranking and rating on Google Play Store

PROJECTS

Generated Pseudo Labels Using Semi-Supervised Learning.

- Preprocessed YouTube user data, generated sentence embeddings using transformer models, and trained on language models such as **BERT** and **RoBERTa**. Achieved up to **83%**(AUPRC). [Github Code Link](#)

Checkpoint Restart.

- Developed checkpointing restart that periodically saves a process's image file onto the desk. When the program crashes, the program could be restarted using a saved checkpoint image. [Github Code Link](#)

PUBLICATIONS

[Johnny Still Can't Opt-out: Assessing the IAB CCPA Compliance Framework](#)

Muhammad Aziz, and Christo Wilson. *Proceedings on Privacy Enhancing Technologies (PoPETS)*, 2024

[Market or Markets? Investigating Google Search's Market Shares Under Vertical Segmentation](#)

Desheng Hu*, Jeffrey Gleason*, [Muhammad Aziz*](#), Nik Guggenberger, and Christo Wilson. *AAI ICWSM*, 2024

SKILLS

- **Languages, Tools, Libraries:** Python, JavaScript, Spark, Node, Pytorch, Pandas, SQL, Go